

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 62-224870

(43)Date of publication of application : 02.10.1987

(51)Int.Cl.

G06K 9/46

G06K 9/34

G06K 9/62

(21)Application number : 61-065640

(71)Applicant : HITACHI LTD

(22)Date of filing : 26.03.1986

(72)Inventor : FUKUDA KOJI  
HINO MASATOSHI  
TABATA KUNIAKI

## (54) DOCUMENT PICTURE PROCESSING SYSTEM

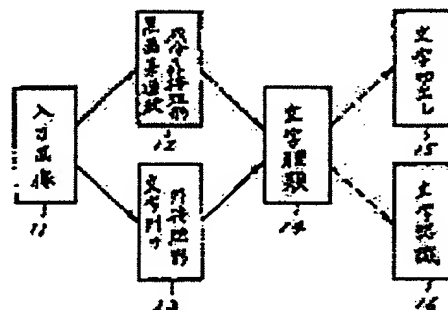
## (57)Abstract:

**PURPOSE:** To automatically discriminate the kind of font at every area when two kinds of character area are mixed in the same document picture by obtaining the frequency distribution as to the position of a circumscribing rectangle of a black picture element continuous area at every character string.

**CONSTITUTION:** For a document picture 11 inputted to a frame memory, the circumscribing rectangle 12 of black picture element continuous components and the circumscribing rectangle 13 of character string 13 are obtained. From the two kinds of circumscribing rectangles 12 and 13 thus obtained out of the input picture 11, the kind of font 14 constituting each character string is discriminated. That is, the frequency distribution as to the position of the circumscribing rectangle of the plural black picture element continuous components that are included in the said rectangle 13, is obtained at every circumscribing rectangle 13 of character string. This frequency distribution varies

depending on whether the character string is a Japanese sentence or an English sentence.

Therefore, from the distribution frequency, the kind of character can be discriminated. In such a way, in case where two kinds of character area are mixed in one document picture, the kinds of character are automatically discriminated at every area.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision  
of rejection]

[Date of requesting appeal against examiner's  
decision of rejection]

[Date of extinction of right]

Copyright (C); 1998,2003 Japan Patent Office

⑨ 日本国特許庁(JP)

⑩ 特許出願公開

⑫ 公開特許公報(A)

昭62-224870

⑤ Int.Cl.<sup>4</sup>

G 06 K 9/46  
9/34  
9/62

識別記号

庁内整理番号

A-6942-5B  
6942-5B  
C-6942-5B

④ 公開 昭和62年(1987)10月2日

審査請求 未請求 発明の数 1 (全9頁)

⑬ 発明の名称 文書画像処理方式

⑭ 特 願 昭61-65640

⑮ 出 願 昭61(1986)3月26日

⑯ 発 明 者 福 田 浩 至 川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

⑰ 発 明 者 樋 野 匡 利 川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

⑱ 発 明 者 田 畑 邦 晃 川崎市麻生区王禅寺1099番地 株式会社日立製作所システム開発研究所内

⑲ 出 願 人 株式会社日立製作所 東京都千代田区神田駿河台4丁目6番地

⑳ 代 理 人 弁理士 小川 勝男 外1名

明 細 書

1. 発明の名称

文書画像処理方式

2. 特許請求の範囲

1. 文字列を含む文書情報を画像データとしてメモリに取り込むステップと、上記画像データから風面連続領域毎の外接矩形を求めるステップと、上記画像データから文字列毎の外接矩形を求めるステップと、文字列外接矩形毎に、それに含まれる複数の風面連続領域外接矩形の位置に関する頻度分布を求めるステップと、上記頻度分布から当該文字列を構成する文字種類を判定するステップとを含むことを特徴とする文書画像処理方式。

2. 第1項記載の文書画像処理方式において、更に、前記文字種類を前提にして当該文字列の少なくとも1部の文字を文字コードに変換するステップを含むことを特徴とする文書画像処理方式。

3. 前記頻度分布が、前記文字列外接矩形の上辺

および下辺に対する前記各風面連続領域外接矩形の離間程度を示すことを特徴とする第1項または第2項記載の文書画像処理方式。

4. 前記頻度分布が、前記風面連続領域外接矩形間の文字列方向の離間程度を示すことを特徴とする第1項または第2項記載の文書画像処理方式。

3. 発明の詳細な説明

〔産業上の利用分野〕

本発明は文書画像の処理方式に関し、更に詳しくは、画像データとして取り込まれた文書情報中の文字認識、あるいは特定の文字領域の切り出しに連した文書画像処理方式に関する。

〔従来の技術〕

光ディスクのような大容量のメモリをファイル装置とした文書画像処理システムにおいては、膨大な文書の検索に必要な索引情報の自動入力化が要望される。文書画像は、本質的に濃淡値をとる写真領域と、文字あるいは図形の如く2値情報で表わされる領域とに大別されるが、索引情報の自動

入力のためには、入力文書画像中の文字領域で特定の文字列、あるいは文字を抽出し、文字認識により照引用語を文字コードに変換する必要がある。

従来、文書画像中の文字を認識する技術として、例えば、アイ・イー・イー・イー、プロシーディング 第6回 アイ・シー・ビー・アール

(IEEE Proceeding 5th ICPR),

1982年、第1023頁～第1026頁において、英文(ローマ字)を対象とし、文字列から互いに接触した位置関係にある文字群(語)を分離して認識する手法が報告されている。また、文字認識のために、文字列から個別文字を分離する技術として、例えば、電子通信学会論文誌, Vol. J67-D, No. 10 (1984年10月)には、日本語文書を対象とした文字切り出しの例が報告されている。しかしながら、これらの報告は、いずれも文書構成文字の種類(英文か和文か)が既知であることを前提としたものであり、文字種類が未知、あるいは複数種類の文字が混在する文書画像を対象としたものではない。

2種類以上の文字が混在する場合、例えば、タイトル、著者、アブストラクト部などの重要部分については日本語の他に英文による記載もなされているような学会論文に対しては適用できない。また、この方式では、全ての黒画素についてストロークの方向、形状等を判定する必要があるため、処理が複雑であり、判定のためのデータ処理に大きなメモリ容量を要するという問題もある。

本発明の目的は、同一文書画像中に2種類の文字領域が混在する場合に、領域毎の文字種類を自動的に判別できるようにした文書画像処理方式を提供することにある。

本発明の他の目的は、特に日本文(和文)領域と英文領域との識別に適した文書画像処理方式を提供することにある。

(問題点を解決するための手段)

上記目的を達成するため、本発明による文書画像処理方式は、文字列を含む文書情報を画像データとしてメモリに取り込むステップと、上記画像データから黒画素連続領域毎の外接矩形を求める

一方、情報処理学会論文誌, Vol. 25, No. 2 (1985年3月)の第257頁～第264頁には、文書構成文字の種類を、文字パターンの視覚的特徴により識別するための方法についての報告がなされている。この報告では、例えば3線素×3線素の小さな単位メッシュで文字を走査し、単位メッシュ内に白画素領域と黒画素領域がただか1個ずつしか現われないようにした場合、単位メッシュのパターン種類が $2^9 = 512$ 個のうちの66個のパターン(合法パターン)に限定されてしまうことに着目し、文字走査時に得られる合法パターンの種類、傾度分布から、文字ストロークの傾斜方向、曲線性、文字の縦横比などの特徴を抽出し、これによりローマ文、ギリシャ文字、ロシア文字、日本文字、ハングル文字等の判別をしている。

(発明が解決しようとする問題点)

然るに、上述した合法パターンを用いる文字種類識別方式は、文書中に現われる合法パターンの統計的な特徴を利用しているため、同一文書中に

ステップと、上記画像データから文字列毎の外接矩形を求めるステップと、文字列外接矩形毎に、それに含まれる複数の黒画素連続領域外接矩形の位置に関する傾度分布を求めるステップと、上記傾度分布から当該文字列を構成する文字種類を判定するステップとを含むことを特徴とする。

(作用)

文書中の文字領域において、各文字のストロークを示す黒画素連続領域毎の外接矩形は、例えば漢字、片仮名、平仮名からなる和文と、大文字、小文字のアルファベットからなる英文とを比較すると、その並び方において差違がある。例えば、英文の場合は、単語と単語との間にスペース部分が存在するが、和文にはない。また、英文では大文字と小文字とで黒画素連続領域の外接矩形の大きさに明らかな違いがあるが、和文の漢字と平仮名、片仮名の間にはアルファベット程の大きさの違いはない。従って、これらの外接矩形の隣接関係、あるいは上下方向でのばらつき方には特徴があり、例えば、文字列外接矩形の上辺および下辺

に対する各黒面素連続領域の外接矩形の離間の程度、あるいは、各黒面素連続領域の外接矩形の文字列方向の離間程度を傾度分布として求めると、分布のパターンから各文字列が和文か英文かを判別できる。同様に、和文とギリシャ語、ロシア語、ドイツ語などの文との区別もできる。

本発明によれば、各文字列毎に、黒面素連続領域の外接矩形の位置に関する傾度分布から、当該文字列の文字種類を判別することにより、例えば、日本語で記載された学会発表文、論文から英文による文字領域を抽出し、これを切り出して抄録を自動的に作成することができる。また、特定の領域について、文字種類を知った上で文字認識処理を実行することができるため、誤認識を少なくして文字コードへの変換を迅速に実行でき、これらの文字コードを索引として自動登録することも可能となる。

#### (実施例)

以下、本発明の実施例について説明する。

第1図は、本発明を実施する文書画像処理シス

く処理する。

すなわち、フレームメモリ7に入力された文書画像11に対して、黒面素連続成分の外接矩形12と各文字列の外接矩形13を求める。黒面素連続成分の外接矩形は、例えば特開昭60-129379号公報等で知られた公知の方法により抽出できる。また、文字列の外接矩形13は、例えば信学技報PRL-83-7において“縦書き横書きと文書からの個別文字切出し法”と題して報告された公知の手法を採用できる。

本発明では、このようにして入力画像から抽出した2種類の外接矩形12と13とから、各文字列を構成する文字種類14を判別し、これに基づいて、文字切り出し処理15、あるいは文字認識処理16を行なう。

第3図(A)は、入力文書の1例として、和文領域20と英文領域21とを含む入力画像11の一部を示し、第3図(B)は、上記入力画像から抽出した文字列矩形31~38と、黒面素連続成分の外接矩形(各文字列外接矩形内の小矩形)

テムのハードウェア構成図であり、1はデータプロセッサ、2は上記プロセッサで必要とするプログラムを格納するためのメモリ、3は文字認識に必要な辞書を格納するためのメモリであり、この例では和文用の辞書と英文用の辞書とが用意される。4は処理結果あるいは中間結果などのデータ類を格納するためのメモリ、5はプロセッサ1に対して各種コマンドあるいは数値データ等を入力するキーボード、6は文書画像の入力装置である。文書画像の入力装置6としては、例えば文書から画像情報を読み取るスキャナが代表的であるが、既にデータが入力されている記録媒体からデータを読取る磁気テープ、磁気ディスク等であってもよい。また、7は入力された文書画像データを一時的に格納するフレームメモリ、8は文書画像データを表示するためのディスプレイ装置、9はハードコピーをとるためのプリンタ装置であり、これらの要素はバス10により接続されている。

本発明では、画像入力装置6から入力された文書画像データをプロセッサ1により、第2図の如

を示す。この例で、文字列の外接矩形(以下、文字列矩形と言う)とそれに含まれる黒面素連続領域の外接矩形(以下、黒面素外接矩形と言う)との関係をみると、和文領域20と英文領域21では次の相違があることが判る。

(1) 和文では、黒面素外接矩形の多くが文字列矩形の底辺に揃っているが、英文の場合は「e」、「j」、「p」、「q」、「y」、「r」など、一部の文字、記号が下側突出ストロークを有し、これによって文字列矩形の底辺が決まるため、黒面素外接矩形の多くは文字列底辺から離れている。

(2) 英文では、ほとんどの文字が連続したストロークからなり、黒面素外接矩形の上辺は、文字列矩形の上辺に一致する第1レベルと、小文字の上辺に相当する第2レベルの2つの顕著なレベルをもつが、和文の場合は、黒面素外接矩形の上辺は、文字列矩形の上辺に集中しており、明確な第2レベルがない。

(3) 各文字列における黒面素外接矩形の隣接関係についてみると、英文の場合は、単語間に大き

な離間部が存在するが、和文には、これがなく、各矩形間の離間距離も不均一である。

本発明は、2種類の文字領域における上述した特徴差に着目して、各文字列毎に構成文字の種類を識別する。

第4図は、複数の風面素外接矩形R1~R7を含む文字列矩形Lについて、各矩形の表示記号を示す。文字列矩形Lの大きさと位置は、左上と右下の風面の座標アドレス(LXmin, LYmin), (LXmax, LYmax)で示され、同様に、各風面素外接矩形の大きさと位置も、それぞれの左上と右下の座標アドレス(RXmin, RYmin), (RXmax, RYmax)で示される。1つの文字列矩形L内に含まれる全ての風面素外接矩形Rは、

$$\begin{aligned} LXmin &\leq RXmin \\ LYmin &\leq RYmin \\ LXmax &\geq RXmax \\ LYmax &\geq RYmax \end{aligned} \quad \dots (1)$$

接矩形領域間の距離Dの分布を求め、第9図の如く、D<sub>1</sub>とD<sub>2</sub>の2つの顕著なピークがある場合は英文、そうでない場合は和文と判断すればよい。

以下、性質(1)、(2)に着目する第1の判定と、性質(3)に着目する第2の判定とを利用した本発明の文書画像処理の手順をプログラム・フローチャートにより説明する。

第10図は、文書画像処理の全体手順を示しており、次の処理ステップからなっている。

処理61: 画像入力装置6から文書画像を入力する。

処理62: 風面素連続成分の外接矩形Rを抽出する。抽出した矩形の座標アドレスを行ごとに第5図の矩形テーブル51に記憶する。ここでは、対角点のアドレスで記憶しているが矩形のアドレスを一意に定義できるテーブル形式ならば例えば、一点の座標アドレスと、矩形の幅、高さを表現する他の表示形式を採用してもよい。

処理63: 文字列矩形Lを抽出して、各アドレスデータを文字列矩形テーブル52に記憶する。

を満足する。入力画像から抽出された上記各矩形の座標値は、例えば第5図に示す如く、データメモリ4に用意した風面素外接矩形テーブル51および文字列矩形テーブル52に、各矩形ID毎に記憶される。

上述した性質(1)、(2)に着目して各文字列の構成文字種類を判別する場合は、例えば第6図に示す如く、各文字列矩形L<sub>j</sub>を高さ方向にn個の領域(この例では領域a~bの8領域)に分割し、この文字列に含まれる風面素外接矩形Rの上辺RXmaxと下辺RYminとが、それぞれどの分割領域に位置しているかを調べ、その頻度分布を求めればよい。RYmaxとRXminの頻度分布は、それぞれ第7図に示す如く、和文と英文で異なった特徴を示すから、例えば上下両端の分割領域aとbに高い頻度を示す文字列は和文、その内側の分割領域bとcに高い頻度をもつ文字列は英文と判断できる。

一方、性質(3)に着目する場合は、第8図に示すように、同じ文字列内に含まれる各風面素外

処理64: 風面素外接矩形テーブル51と文字列矩形テーブル52を参照して文字列が和文であるか英文であるかを判定する第1の判定処理を行なう。

処理65: 処理64で判定できなかった文字列に対して、和文か英文かの第2の判定処理を行なう。

処理66: 上記判定結果を参照して、辞書メモリ3を利用して文字認識、文字切出し等の処理を行なう。

第11図は第1の判定処理64の詳細を示すフローチャートであり、次のステップからなっている。

処理701: 各文字列毎に風面素外接矩形Rの上辺分布頻度をカウントするためのカウンタUCと、下辺分布頻度をカウントするためのカウンタDCをそれぞれクリアする。

処理702: 風面素外接矩形テーブル51から矩形ID順に矩形Rの座標データを読み込む。

処理703: テーブル52を参照し、矩形Rが

含まれる文字列 $L_j$ を検索する。包含条件は前述した式(1)による。

処理704: 文字列 $L_j$ を高さ方向に $n$ 個に分割する。各分割領域の上限アドレス $A_{\max}(i)$ と下限アドレス $A_{\min}(i)$ は次式で定義される。

$$A_{\max}(i) = LY_{\min} + (LY_{\max} - LY_{\min}) \cdot i / n$$

$$A_{\min}(i) = LY_{\min} + (LY_{\max} - LY_{\min}) \cdot (i-1) / n$$

処理705: 矩形 $R$ の上辺アドレス $RY_{\max}$ が上記分割領域のどれに含まれるかを調べ、該当するカウンタ $UC(i)$ をインクリメントする。分割領域のインデックス $i$ と $RY_{\max}$ との関係は、

$$A_{\min}(i) \leq RY_{\max} \leq A_{\max}(i)$$

により求める。

処理706: 矩形 $R$ の底辺アドレス $RY_{\min}$ が上記分割領域のどれに含まれるかを調べ、該当するカウンタ $DC(i)$ をインクリメントする。分

が共に満足するかどうかを調べる。もし、満足すれば、この文字列は和文と判定できる。

処理709: 当該文字列が和文であることを示すため、例えばテーブル52の判定結果を示す欄 $F$ に和文表示コードを記録する。

処理710: 処理708で和文と判定できなかった場合、カウンタ $UC$ の最大カウント値が最上端の分割領域 $h$ 以外であり、且つ、カウンタ $DC$ の最大カウント値が最下端の分割領域 $a$ 以下に生じていれば、この文字列は英文であると判定する。すなわち、

$$\max\{UC(i)\} \neq UC(h), (i = a \sim h)$$

$$\max\{DC(i)\} \neq DC(a), (i = a \sim h)$$

が共に満足する場合、英文と判断して処理ステップ711に進み、そうでない場合はステップ712に進む。

処理711: 文字列矩形テーブル52の下欄に英文であることを示すコードを記録する。

割領域のインデックスと $RY_{\min}$ との関係は、

$$A_{\min}(i) \leq RY_{\min} \leq A_{\max}(i)$$

により求める。

これらの処理により、矩形 $R$ の上辺を下辺が、第6図の分割領域 $a \sim h$ のいずれに位置しているかが、各文字列毎のカウンタ $UC$ 、 $DC$ に頻度として記憶される。

処理707: テーブル51の全ての矩形について判定処理を終えたかどうかをチェックし、NOならば処理702に戻り、YESなら処理708に進む。

処理708: 各文字列毎のカウンタ $UC$ 、 $DC$ を参照し、それぞれの最大カウント値が文字列矩形の上端または下端の分割領域に対応するカウンタで生じているかどうかを判定する。すなわち、

$$\max\{UC(i)\} = UC(h), (i = a \sim h)$$

$$\max\{DC(i)\} = DC(a), (i = a \sim h)$$

処理712: 文字列矩形テーブル52の下欄に判定不能である旨の表示記号を記録する。

処理713: 全ての文字列について文字種類判定を終了したかどうかを判定し、未処理の文字列が残っている場合は処理ステップ708に戻り、そうでない場合は、このルーチンを終了する。

第12図は第2の判定処理65の詳細を示すフローチャートである。

処理800: 文字列テーブル52の下欄をチェックし、前回の第1の判定で判定不能と判断された文字列を見つける。もし、なければ、このルーチンを終了し、あれば次のステップ801に進む。

処理801: 文字列矩形 $L_j$ データを読み取る。

処理802: 文字列矩形 $L_j$ に含まれる全ての黒画素外接矩形を検索する。この場合の検索条件は式(1)であり、既に行なった第1の判定処理64で各矩形と $L_j$ との対応関係が保存(例えばテーブル51に対応する $L_j$ の矩形IDを記憶)してあれば、それを利用してもよい。

処理803: 隣接する黒画素外接矩形間の距離

Dの分布をカウントするためのカウンタLCを初期化する。

処理804: ステップ802で検索された黒面素外接矩形の1つR<sub>i</sub>について、隣接矩形R<sub>i+1</sub>とのX方向の離間距離DをそれぞれのX座標を用いて求める。

処理805: 距離Dに対応するカウンタLC(n)をインクリメントする。カウンタのインデックスnは、距離Dを所定の値ずつ離散化することにより求まる。

処理806: 文字列矩形L<sub>j</sub>に含まれる全ての黒面素外接矩形について、上記距離判定を終了したか否かをチェックし、未処理矩形があれば、処理ステップ804に戻り、そうでなければ処理ステップ807に進む。

処理807: カウンタLCの値をチェックし、第9図の如く、ピーク値が互いに離間した2箇所にあるか否かを調べる。もし、あれば英文と判定して処理ステップ808に進み、そうでなければ処理ステップ809に進む。ピーク値の判定は、

する。

処理811: 全文字列について処理が終了したか否かを判定し、未処理の文字列があれば処理ステップ800に戻り、なければこのルーチンを終了する。

以上、判定処理64と65について説明したが、これらの処理は互いに入れ換えてもよく、また、いずれか一方のみを用いるようにしてもよい。また、判定処理64あるいは65において、文字種別を判定できない場合、前後の文字列の判定結果から不明文字列の種別を推定するようにしてもよい。

尚、第3図の英文領域21を参照すると、アルファベットの大多数は、各文字のストロークが連続しており、黒面素外接矩形領域が縦方向に分離されるケース(例えば小文字の「l」、「j」)は極めて稀である。即ち、英文の場合は、RYmin、RYmaxのそれぞれの値の分散が小さいため、これを利用して和文との区別をつけることができる。例えば、1つの文字列に含まれる黒面素外接矩形

ピーク点を示す所定の傾きKに対して、

$$LC(i) - LC(i-1) > K$$

$$LC(i) - LC(i+1) > K$$

を満たすカウンタLC(i)が2つ(i = i<sub>1</sub>, i<sub>2</sub>)存在し、

$$|i_1 - i_2| > p$$

但し、pは所定の距離を示す値か否かを判定すればよい。

処理808: 文字列矩形L<sub>j</sub>が英文であることを示すコードを、テーブル52の欄Fに記録する。

処理809: カウンタLCの各カウント値の分散を所定の閾値と比較し、これより大きい場合は処理ステップ810に進み、そうでない場合は処理ステップ811に進む。

処理810: 文字列矩形L<sub>j</sub>を和文と判断し、テーブル52の欄Fにその旨の表示コードを記録

の数をmとすると、RYminの分散Sminは、

$$S_{min} = \frac{\sqrt{\sum_{i=1}^m (RYmin(i) - \overline{RYmin})^2}}{m}$$

で求め、RYmaxの分散Smaxも同様にして求める。従って、分散の閾値Smaxφ、Sminφに対して、和文は

$$\begin{aligned} S_{min} &> S_{min\phi} \\ \text{and} \quad S_{max} &> S_{max\phi} \end{aligned}$$

で判断できる。尚、Smaxφ、Sminφの値は、文字列の高さの関数として定義することにより、文字の大きさにより変えることができる。

—この判定は、第1、第2の判定に代え、あるいは、これらを初う第3の判定として、第10図のフローチャートに加えることができる。

以上の方法により、文字列毎の文字種別が自動的に判別できるため、第2図に示したように、こ



これらの文字種別を前提として、文字切出し、あるいは文字認識を行なうことができる。

(発明の効果)

本発明によれば、文書画像データ中の文字領域において、構成文字の種類を文字列毎に自動的に判別できるため、この判別結果を利用して、特定文字領域の抽出(切り出し)、文字認識等を行なうことができ、その結果を画像ファイルの索引、あるいは抄録等の補助情報として自動的に登録するような文書画像処理システムにおいて有効である。

#### 4. 図面の簡単な説明

第1図は本発明を実施するための文書画像処理システムの1例を示す全体構成図、第2図は本発明による文書画像処理の概略手順を説明するための図、第3図(A)、(B)はそれぞれ入力文書画像と、これから抽出された黒画素連続領域の外接矩形および文字列外接矩形の具体例を示す図、第4図は文字列外接矩形と黒画素連続領域外接矩形の表示データについての説明図、第5図は上記各

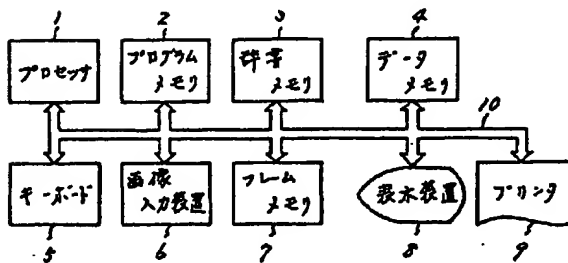
矩形の格納テーブルの構成を示す図、第6図および第7図は文字種類の第1の判定方法における頻度分布のカウントと頻度分布の特性を説明するための図、第8図と第9図は文字種類の第2の判定方法における領域分布のカウントと頻度分布の特性を説明するための図、第10図は本発明による画像処理の全体フローチャート、第11図は第1の判定方法を実施するためのフローチャート、第12図は第2の判定方法を実施するためのフローチャートである。

#### 5. 符号の説明

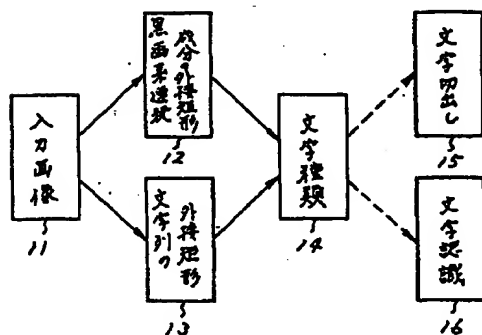
- |                 |             |
|-----------------|-------------|
| 1: プロセッサ        | 2: プログラムメモリ |
| 3: 辞書メモリ        | 4: データメモリ   |
| 5: キーボード        | 6: 画像入力装置   |
| 7: フレームメモリ      | 8: 表示装置     |
| 20: 和文領域        | 21: 英文領域    |
| L: 文字列外接矩形      |             |
| R: 黒画素連続領域の外接矩形 |             |

代理人 弁理士 小川勝男

第1図



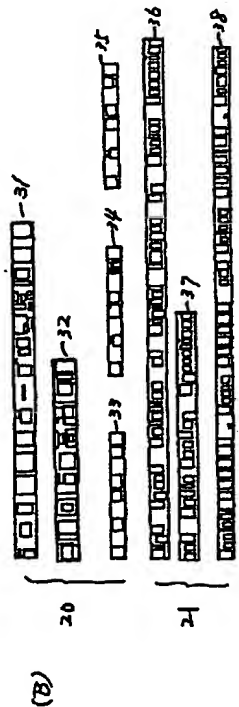
第2図



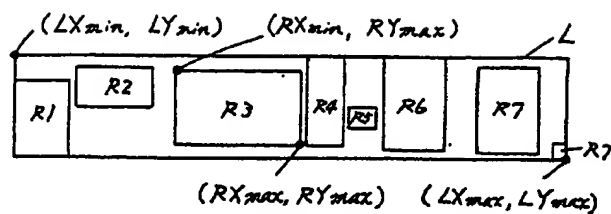
第3図

ラスタ走査とテーブル参照による  
画像回転の高速処理

High Speed Rotation of Digital Images by Raster Scanning  
and Table-Lookup Operations  
Kunishi TABATA, Haruo TAKEDA and Teisuo MACHIDA, Members



第4図

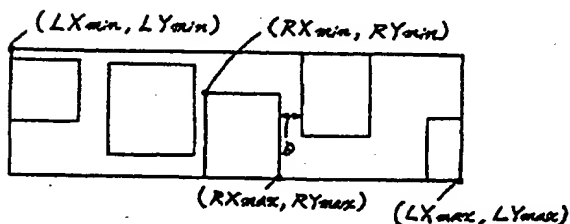


第5図

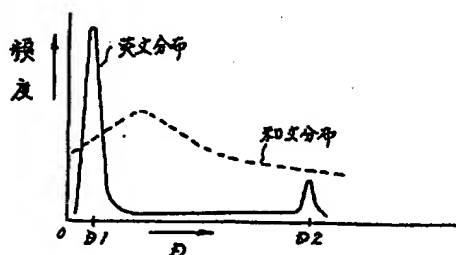
矩形ID	RXmin	RYmin	RXmax	RYmax
R1	30	30	35	35
R2	38	28	50	41
R3	55	32	73	
R4	78	29		
...	...	...	...	...

矩形ID	LXmin	LYmin	LXmax	LYmax	F
L1	30	28	354	45	
L2	25	53	360	76	
L3	25	88	27		
L4	24	111			
...	...	...	...	...	...

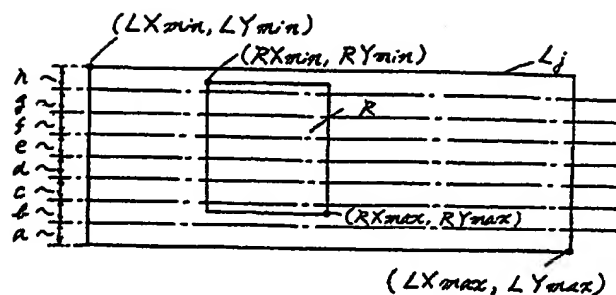
第8図



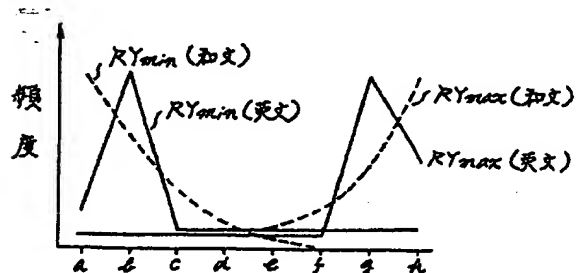
第9図



第6図



第7図



第10図

